# BMAAR Assignment 1

1. Write a note on uniform and normal distribution.

- Normal Distribution is a probability distribution which peaks out in the middle and gradually decreases towards both ends of axis. It is also known as gaussian distribution and bell curve because of its bell like shape.
- Uniform Distribution is a probability distribution where probability of x is constant. That is to say, all points in range are equally likely to occur consequently it looks like a rectangle. Formula for Uniform probability distribution is f(x) = 1/(b-a), where range of distribution is [a, b].

### Generation of Random Numbers with uniform and normal distribution in R.

1. Illustration of sample, runif and rnorm functions with examples

```r
cat("Sample Function Output:",sample(x=10,size = 9,replace = TRUE),"\n");
```

```
## Sample Function Output: 4 1 4 8 7 4 8 8 2
```

```r
cat("Runif Function Output:",runif(n = 9,min = 10,max = 100),"\n");
```

```
## Runif Function Output: 83.18713 75.47574 96.30052 94.8273 86.01706 46.609 22.1394 50.10096 74.92133
```

```r
cat("Rnorm Function Output:",rnorm(n = 9,mean = 0, sd = 1))
```

```
## Rnorm Function Output: -1.353557 -1.135841 -0.2881581 -0.3120521 0.3791908 0.407889 -0.1576672 1.7003
```

2. Write a R program to create a list of random numbers in normal distribution and count occurrences of each value.

```r
random <- rnorm(n = 10,mean = 0,sd = 1)
print(table(random))
```

```
## random
## -0.990609323326192 -0.937098596747361 -0.754414455997871 -0.552845926718838
##                  1                  1                  1                  1
##  0.289958709779913  0.298139636701745  0.404862836323234  0.927008468817732
##                  1                  1                  1                  1
##   1.06416401597498   2.66292026609175
##                  1                  1
```

3. Write a R program to create a vector which contains 10 random integer values between -50 and +50

```r
random <- runif(10,min = -50,max = 50)
print(random)
```

```
##  [1]  34.553177  44.578858  -9.578962 -27.905623  33.814746 -34.636876
##  [7] -37.350264 -10.063206 -31.062423 -23.692424
```

4. Use the sample function to obtain a random sample of 10 realizations in a biased coin experiment

```r
sample(c('H','T'),10,replace = TRUE)
```

```
##  [1] "T" "T" "T" "T" "H" "T" "H" "T" "H" "H"
```

5. Create a fair dice (with possible outcomes from 1 to 6) and determine the arithmetic mean and standard deviation of throwing it 10,000 times.

```
dice <- c(1:6);
throws <- sample(dice,10000,replace=TRUE);
cat("Mean:",mean(throws),"\n");
```

```
## Mean: 3.4977
```

```
cat("Standard Deviation:",sd(throws))
```

```
## Standard Deviation: 1.713268
```

6. The most popular German lottery is known as 6 aus 49, in which a total of 7 numbers are randomly drawn: First, 6 unique numbers are randomly drawn out of the numbers from 1 to 49. Second, a single-digit "Superzahl" between 0 and 9. Simulate this lottery and run it once.

```
first_draw <- sample(c(1:49),6);
second_draw <- sample(c(0,9),1);
cat("First Draw =>",first_draw,"\n");
```

```
## First Draw => 13 11 3 45 30 25
```

```
cat("Second Draw =>",second_draw)
```

```
## Second Draw => 0
```

7. Suppose we select a SRS of n = 3 balls from an urn containing a population of N = 6 balls (painted with the numbers 1, 2, 3, 4, 5, and 6). List the sample space S of all possible outcomes.

```
# install.packages("combinat")
library("combinat")
```

```
##
## Attaching package: 'combinat'
```

```
## The following object is masked from 'package:utils':
##
##     combn
```

```
sample_space <- combinat::combn(6,3)
print(sample_space)
```

```
##      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12] [,13] [,14]
## [1,]    1    1    1    1    1    1    1    1    1     1     2     2     2     2
## [2,]    2    2    2    2    3    3    3    4    4     5     3     3     3     4
## [3,]    3    4    5    6    4    5    6    5    6     6     4     5     6     5
##      [,15] [,16] [,17] [,18] [,19] [,20]
## [1,]     2     2     3     3     3     4
## [2,]     4     5     4     4     5     5
## [3,]     6     6     5     6     6     6
```

## Write a note on Probability Distributions and its types.

- Common probability distributions include the binomial distribution, Poisson distribution, and uniform distribution. Certain types of probability distributions are used in hypothesis testing, including the standard normal distribution, the F distribution, and Student's t distribution.
- Discrete Probability Distribution A discrete distribution describes the probability of occurrence of each value of a discrete random variable. The number of spoiled apples out of 6 in your refrigerator can be

an example of a discrete probability distribution. Each possible value of the discrete random variable can be associated with a non-zero probability in a discrete probability distribution.

- Binomial Distribution The binomial distribution is a discrete distribution with a finite number of possibilities. When observing a series of what are known as Bernoulli trials, the binomial distribution emerges. A Bernoulli trial is a scientific experiment with only two outcomes: success or failure. Consider a random experiment in which you toss a biased coin six times with a 0.4 chance of getting head. If 'getting a head' is considered a 'success', the binomial distribution will show the probability of r successes for each value of r. The binomial random variable represents the number of successes (r) in n consecutive independent Bernoulli trials.

- Poisson Distribution A Poisson distribution is a probability distribution used in statistics to show how many times an event is likely to happen over a given period of time. To put it another way, it's a count distribution. Poisson distributions are frequently used to comprehend independent events at a constant rate over a given time interval. Siméon Denis Poisson, a French mathematician, was the inspiration for the name.

- Continuous Probability Distributions A continuous distribution describes the probabilities of a continuous random variable's possible values. A continuous random variable has an infinite and uncountable set of possible values (known as the range). The mapping of time can be considered as an example of the continuous probability distribution. It can be from 1 second to 1 billion seconds, and so on. The area under the curve of a continuous random variable's PDF is used to calculate its probability. As a result, only value ranges can have a non-zero probability. A continuous random variable's probability of equaling some value is always zero.

- Normal Distribution Normal Distribution is one of the most basic continuous distribution types. Gaussian distribution is another name for it. Around its mean value, this probability distribution is symmetrical. It also demonstrates that data close to the mean occurs more frequently than data far from it. Here, the mean is 0, and the variance is a finite value.

- Continuous Uniform Distribution In continuous uniform distribution, all outcomes are equally possible. Each variable has the same chance of being hit as a result. Random variables are spaced evenly in this symmetric probabilistic distribution, with a 1/ (b-a) probability.

#### Probability Distributions: Demonstration of CDF and PDF uniform and normal, binomial & Poisson distributions in R.

1. Illustration of PDF & CDF functions of normal, binomial & Poisson distributions.

```
cdf <- pnorm(3,0,1)
cdf
```

```
## [1] 0.9986501
```

```
pdf <- dnorm(3,0,1)
pdf
```

```
## [1] 0.004431848
```

2. Calculate the following probabilities:

- Probability that a normal random variable with mean 22 and variance 25
    1. lies between 16.2 and 27.5

```
mean_1 <- 22
std <- sqrt(25)
pnorm(27.5,mean = mean_1,sd = std) - pnorm(16.2,mean = mean_1,sd = std)
```

```
## [1] 0.7413095
```

```
2. is greater than 29
```

```r
1 - pnorm(29,mean = mean_1,sd = std)
```

```
## [1] 0.08075666
```

```
3. is less than 17
```

```r
pnorm(17,mean = mean_1,sd = std)
```

```
## [1] 0.1586553
```

```
4. is less than 15 or greater than 25
```

```r
pnorm(15,mean = mean_1,sd = std) + (1- pnorm(25,mean = mean_1,sd = std))
```

```
## [1] 0.3550098
```

- Probability that in 60 tosses of a fair coin the head comes up
    1. 20,25 or 30 times

```r
dbinom(20, size=60, prob=0.5) + dbinom(25, size=60, prob=0.5) + dbinom(30, size=60, prob=0.5)
```

```
## [1] 0.1512435
```

```
2. less than 20 times
```

```r
pbinom(20,size = 60,prob = 0.5)
```

```
## [1] 0.006744647
```

```
3. between 20 and 30 times
```

```r
pbinom(30, size=60, prob=0.5) - pbinom(20, size=60, prob=0.5)
```

```
## [1] 0.5445444
```

- A random variable X has Poisson distribution with mean 7. Find the probability that
    1. X is less than 5

```r
ppois(5,lambda = 7)
```

```
## [1] 0.3007083
```

```
2. X is greater than 10 (strictly)
```

```r
1 - ppois(10, lambda=7)
```

```
## [1] 0.09852079
```

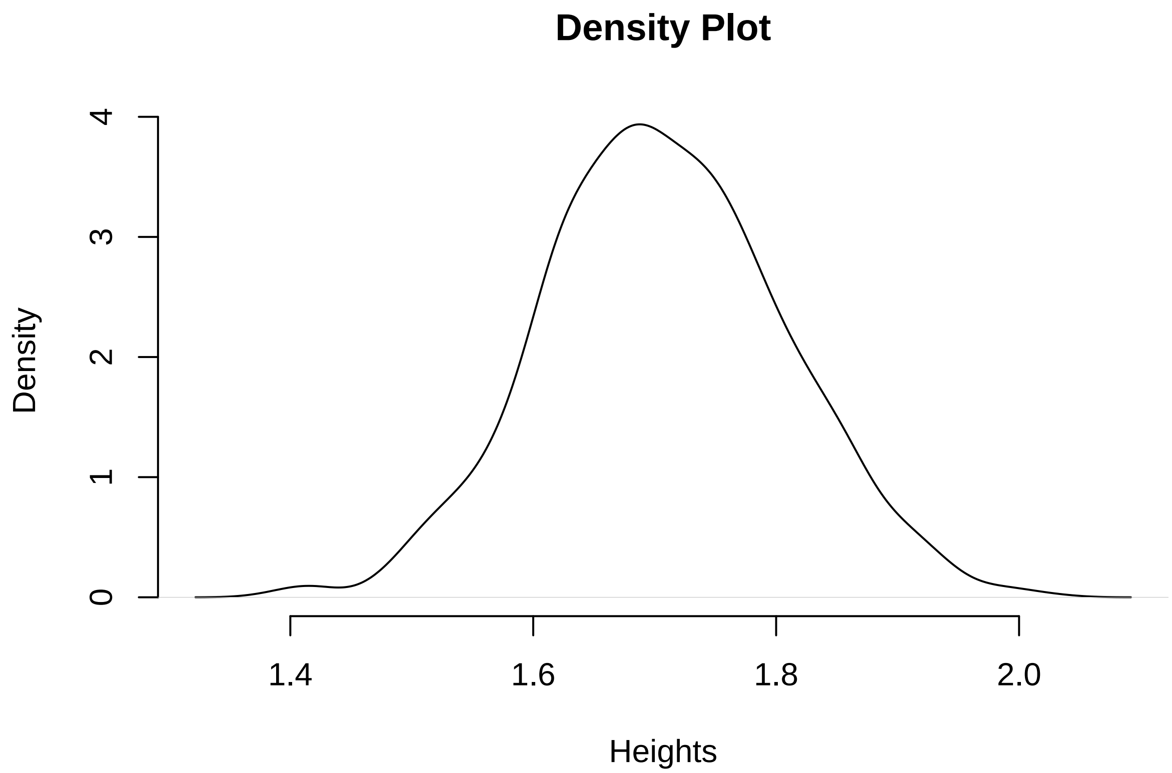
```
3. X is between 4 and 16
```

```r
ppois(16, lambda=7) - ppois(4, lambda=7)
```

```
## [1] 0.8260502
```

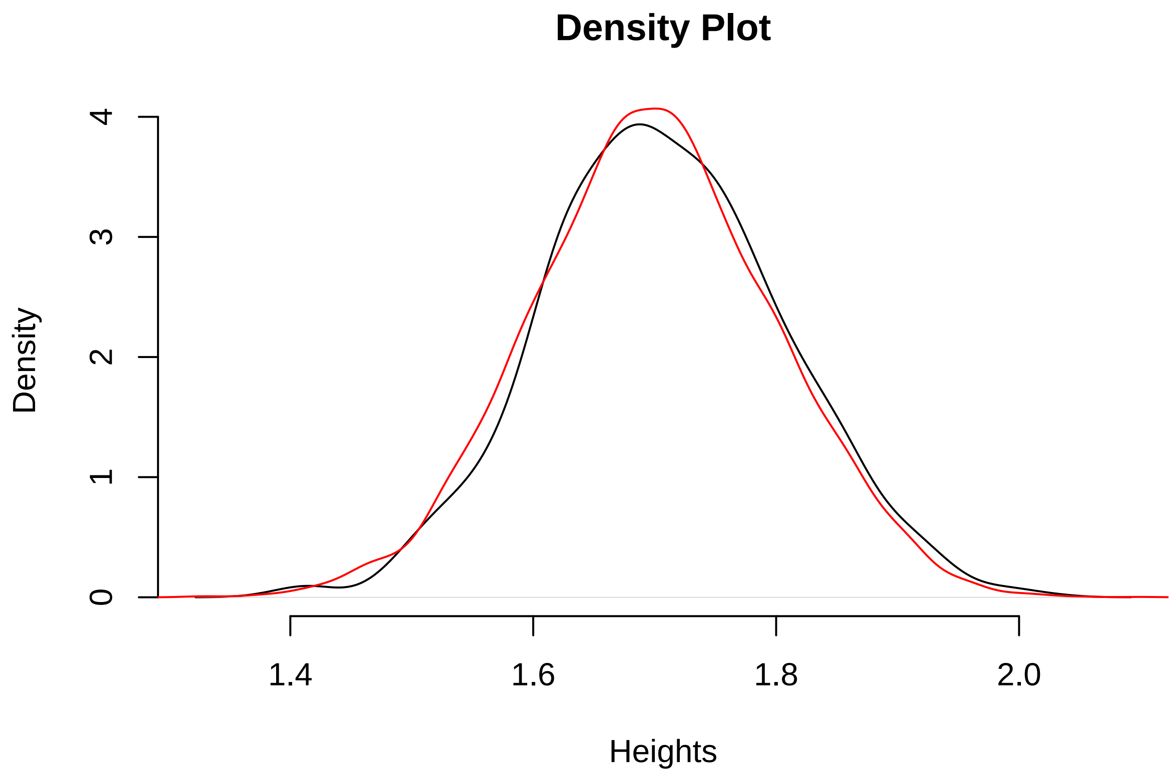
3. Simulate normal distribution values. Imagine a population in which the average height is 1.70 m with a standard deviation of 0.1. Use rnorm to simulate the height of 1000 people and save it in an object called heights.
    1. Plot the density of the simulated values.

```r
library(ggplot2)
heights <- rnorm(1000,1.70,0.1)
# hist(heights,xlab = "Height")
plot(density(heights),frame=FALSE,main = "Density Plot",xlab = "Heights")
```




2. Generate 10000 values with the same parameters and plot the respective density function on top of the previous plot in red to differentiate it.

```
heights_2 <- rnorm(10000,1.70,0.1)
plot(density(heights),frame=FALSE,main = "Density Plot",xlab = "Heights")
lines(density(heights_2),col="red")
```




4. You roll a die 100 times and get just 10 sixes?
    1. What is the probability of getting just 10 sixes?

```
dbinom(10,100,1/6)
```

```
## [1] 0.02140327
```

2. What is the probability of getting 10 or fewer sixes?

```
pbinom(10,100,1/6)
```

```
## [1] 0.04269568
```

3. Draw the probability distribution.

```
rolls <- rbinom(1,100,1/6)
rolls
```

```
## [1] 10
```

4. Simulate the described experiment 1000 times and compute the empirical distribution. Compare it to the theoretical one. Then do the same with 1,000,000 simulations.

```
# ecdf()
```

5. Using the function rbinom to generate 10 unfair coin tosses with probability success of 0.3. Set the seed to 1.

```
set.seed(1)
rbinom(10,1,0.3)
```

```
##  [1] 0 0 0 1 0 1 1 0 0 0
```

6. Simulate normal distribution values. Imagine a population in which the average height is 1.70 m with an standard deviation of 0.1, using rnorm simulate the height of 100 people and save it in an object called heights.

```r
heights <- rnorm(100,1.70,0.1)
```

```
1. What’s the probability that a person will be smaller or equal to 1.90 m ?
```

```r
pnorm(1.90,1.70,0.1)
```

```
## [1] 0.9772499
```

```
2. What’s the probability that a person will be taller or equal to 1.60 m?
```

```r
1 - pnorm(1.60,1.70,0.1)
```

```
## [1] 0.8413447
```